

Clean Copy of Amended Claims

---

1 1. An automated method for setting up a natural language interface in a  
2 Web site comprising the steps of:  
3 defining a hierarchy of topics into which individual documents or  
4 Web pages can be classified;  
5 generating a keyword index for those documents; and  
6 for each topic in the hierarchy, associating a set of n-grams to a  
7 topic in the topic hierarchy, which set of n-grams is distinctive to that topic  
8 and wherein the n-grams may be sparse n-grams or non-sparse n-grams.

1 2. The automated method for setting up a natural language interface in a  
2 Web site recited in claim 1, wherein the step of generating a keyword  
3 index comprises the step of extracting sparse n-grams of keywords for each  
4 group of pages in the topic hierarchy.

1 3. The automated method for setting up a natural language interface in a  
2 Web site recited in claim 1, further comprising the step of optionally  
3 reviewing and editing the keyword index.

1 4. An automated method for setting up a natural language interface in a  
2 Web site comprising the steps of:  
3 automatically inducing a topic hierarchy by examining a structure  
4 of the Web site;  
5 creating n-grams from pages in the Web site that are associated  
6 with a topic in the topic hierarchy and wherein the n-grams may be sparse  
7 n-grams or non-sparse n-grams; and  
8 creating rules from the n-grams, wherein each topic has associated  
9 rules that are used to decide if a new input document or query references  
10 the topic.

- ab
- 1 5. The automated method for setting up a natural language interface in a
  - 2 Web site recited in claim 4, wherein the step of creating rules is performed
  - 3 automatically and further comprising the optional step of manually editing
  - 4 the rules.

New claim 6 is as follows:

- 1 6. The automated method for setting up a natural language interface in a
  - 2 Web site recited in claim 1, further comprising the step of converting the
  - 3 set of n-grams to classification rules.
- 
-